

VOICE OUTPUT/INPUT DEVICE, VOICE OUTPUT/INPUT METHOD AND VOICE OUTPUT/INPUT PROGRAM

Publication number: JP2003248498

Publication date: 2003-09-05

Inventor: NIIHARA YASUYUKI; YAMAGUCHI YOSHIKAZU; TAKAHASHI SATOSHI;
IMAMURA AKIHIRO; KAWABATA TAKESHI

Applicant: NIPPON TELEGRAPH & TELEPHONE

Classification:

- international: G10L11/02; G10L15/00; G10L15/04; G10L15/22; G10L11/00; G10L15/00;
(IPC1-7): G10L15/22; G10L11/02; G10L15/00; G10L15/04

- european:

Application number: JP20020048650 20020225

Priority number(s): JP20020048650 20020225

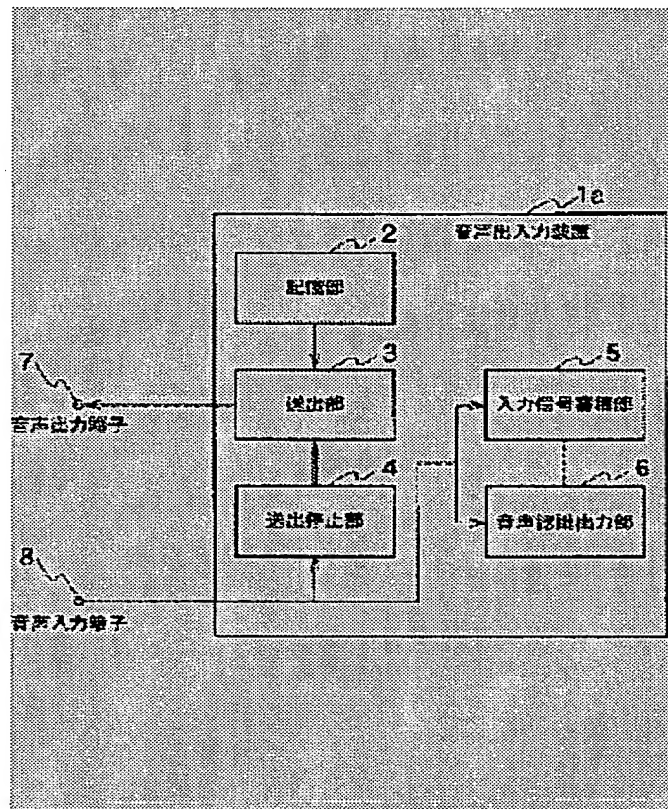
Report a data error here

Abstract of JP2003248498

PROBLEM TO BE SOLVED: To provide a voice output/input device, a voice output/input method and a voice output/input program, by which transmission of aural signals such as voice guidance is stopped only when voice is inputted.

SOLUTION: The voice output/input device 1a is provided with a storage section 2 which stores voice signals, a transmitting section 3 which transmits stored voice signals, a transmission stopping section 4 which discriminates to determine whether the signals inputted from a receiving site of the aural signals during the transmission of the aural signals by the section 3 are voice or not based on the sound pressure level and the fundamental frequency of the inputted signals and stops the transmission of the aural signals when the inputted signals are voice, an input signal accumulating section 5 which accumulates the inputted signals and a voice recognition output section 6 which recognizes the discriminated voice and outputs the voice.

COPYRIGHT: (C)2003,JPO



RESULT LIST

1 result found in the Worldwide database for:

JP2003248498 (priority or application number or publication number)

(Results are sorted by date of upload in database)

**1 VOICE OUTPUT/INPUT DEVICE, VOICE OUTPUT/INPUT METHOD AND VOICE
OUTPUT/INPUT PROGRAM**

Inventor: NIIHARA YASUYUKI; YAMAGUCHI YOSHIKAZU; (+3)

Applicant: NIPPON TELEGRAPH & TELEPHONE

EC:

IPC: **G10L11/02; G10L15/00; G10L15/04** (+7)

Publication info: **JP2003248498** - 2003-09-05

Data supplied from the *esp@cenet* database - Worldwide

Partial Translation of JP 2003-248498

Publication Date: September 5, 2003

Application No.: 2002-48650

Filing Date: February 25, 2002

Applicant: NIPPON TELEGRAPH & TELEPHONE

Inventor: Yasuyuki NIIHRA

Inventor: Yoshikazu YAMAGUCHI

Inventor: Satoshi TAKAHASHI

Inventor: Akihiro IMAMURA

Inventor: Takeshi KAWABATA

[0023]

The transmission stopping section 4 stops transmission of a voice guidance if a signal is inputted during the transmission of the voice guidance by the transmitting section 3 via a sound input terminal 8 from the microphone or the like, and if the signal is a sound. In the invention, the sound indicates a voice or a like component which is uttered in response to the voice guidance.

[0024]

In view of the above, the transmission stopping section 4 includes, as shown in FIG. 2, a judger 41 (corresponding to a judging means in the claims) for judging whether the signal (hereinafter, called as "input signal") inputted to the transmission stopping section 4 is a sound, and a stop signal transmitter 45.

[0025]

The judger 41 has a measurement calculator 42 corresponding to a measurement calculating means in the claims, a comparator/analyzer 43

corresponding to a comparing/analyzing means in the claims, and a judgment executer 44 corresponding to a judgment executing means in the claims.

[0026]

The measurement calculator 42 measures a sound pressure level of the input signal, and calculates a fundamental frequency thereof. The comparator/analyzer 43 compares the sound pressure level measured by the measurement calculator 42 with a predetermined threshold value, and analyzes the fundamental frequency calculated by the measurement calculator 42.

[0027]

The judgment executer 44 judges that the input signal is a sound, as a result of the comparison and the analysis by the comparator/analyzer 43, if the sound pressure level of the input signal is larger than the predetermined threshold value, and if a pitch component as information representing that the input signal is a sound is detected from the fundamental frequency.

[0028]

Then, the stop signal transmitter 45 transmits, to the transmitting section 3, a transmission stop signal indicating stop of transmitting the voice guidance, based on the judgment by the judgment executer 44. The transmitting section 3 stops the transmission of the voice guidance upon receiving the transmission stop signal.

[0029]

If, on the other hand, the sound pressure level is smaller than the predetermined threshold value, or if the pitch component is not included, the transmission of the voice guidance is continued instead of stop.

[0030]

The transmission of the voice guidance is stopped until the sound input is completed. Upon detecting that the sound input is completed, the transmitting section 3 resumes the transmission of the voice guidance.

[0031]

The predetermined threshold value is settable to an arbitrary value. The fundamental frequency depends on a vibration frequency of vocal cords of a speaker. Generally, a male speaker has a vibration frequency in the range from 50 to 150 Hz, and a female speaker has a vibration frequency in the range from 200 to 300 Hz.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-248498

(P2003-248498A)

(43) 公開日 平成15年9月5日 (2003.9.5)

(51) Int.Cl. ⁷	識別記号	F I	テームト* (参考)
G 1 0 L 15/22		G 1 0 L 3/00	5 7 1 T 5 D 0 1 5
11/02			5 1 3 B
15/00			5 1 3 A
15/04			B
			5 5 1 A
審査請求 未請求 請求項の数 6 O L (全 8 頁)			

(21) 出願番号 特願2002-48650 (P2002-48650)

(22) 出願日 平成14年2月25日 (2002.2.25)

(71) 出願人 000004226

日本電信電話株式会社

東京都千代田区大手町二丁目3番1号

(72) 発明者 新原 康之

東京都千代田区大手町二丁目3番1号 日

本電信電話株式会社内

(72) 発明者 山口 義和

東京都千代田区大手町二丁目3番1号 日

本電信電話株式会社内

(74) 代理人 100083806

弁理士 三好 秀和 (外1名)

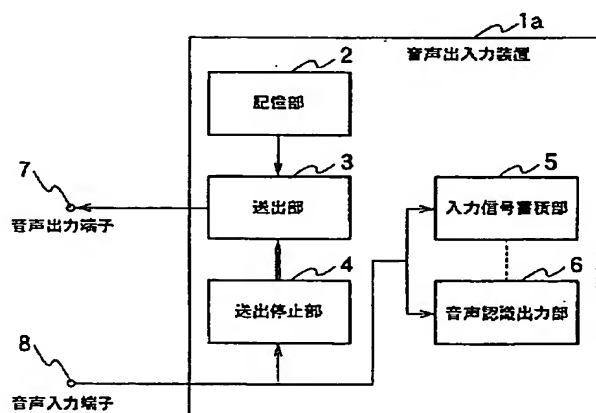
最終頁に続く

(54) 【発明の名称】 音声入出力装置、音声入出力方法、及び音声入出力プログラム

(57) 【要約】

【課題】 音声が入力された時にのみ音声ガイダンス等の音声信号の送出を停止する音声入出力装置、音声入出力方法、及び音声入出力プログラムを提供する。

【解決手段】 音声信号を記憶する記憶部2と、記憶された音声信号を送出する送出部3と、送出部3による音声信号の送出中に音声信号の受信前から入力された信号が音声であるか否かを入力信号の音圧レベル及び基本周波数をもとに判定し、入力された信号が音声であるときには音声信号の送出を停止させる送出停止部4と、入力信号を蓄積する入力信号蓄積部5と、判定がなされた音声を確認し、出力する音声認識出力部6とを有する音声入出力装置1aを設ける。



【特許請求の範囲】

【請求項 1】 音声信号を送出する送出手段と、
前記送出手段による音声信号の送出中に該音声信号の受信先から入力された信号が音声であるか否かを判定し、
該入力された信号が該音声であるときには前記音声信号の送出を停止させる送出停止手段とを有することを特徴とする音声出力装置。

【請求項 2】 前記判定手段は、
前記入力された信号の音圧レベルの測定及び該入力された信号の基本周波数の算出を行う測定算出手段と、
前記測定された音圧レベルと所定の閾値との比較及び前記算出された基本周波数の解析を行う比較解析手段と、
前記測定した音圧レベルが前記閾値より大であり、且つ前記算出された基本周波数から前記入力された信号が音声であることを示す情報が検出された場合、該入力された信号が音声であると判定する判定実行手段とを有することを特徴とする請求項 1 に記載の音声出力装置。

【請求項 3】 前記入力された信号を蓄積する入力信号蓄積手段を有することを特徴とする請求項 1 又は 2 に記載の音声出力装置。

【請求項 4】 前記判定がなされた音声を認識し、出力する音声認識出力手段を有することを特徴とする請求項 1 から 3 のいずれかに 1 項に記載の音声出力装置。

【請求項 5】 音声信号を送出する送出ステップと、
前記送出ステップにおける音声信号の送出中に該音声信号の受信先から入力された信号が音声であるか否かを判定し、
該入力された信号が該音声であるときには前記音声信号の送出を停止させる送出停止ステップとを有することを特徴とする音声出力方法。

【請求項 6】 音声信号を送出する送出ステップと、
前記送出ステップにおける音声信号の送出中に該音声信号の受信先から入力された信号が音声であるか否かを判定し、
該入力された信号が該音声であるときには前記音声信号の送出を停止させる送出停止ステップとを実行させることを特徴とする音声出力プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、通信回線を通じて通信端末との間で音声信号の出力を行う音声出力装置、音声出力方法、及び音声出力プログラムに関する。

【0002】

【従来の技術】通信端末に向けて音声信号である音声ガイダンスを送出し、この音声ガイダンスの受信先からの入力信号を受け付ける音声出力装置は、アンケート調査や苦情受付をはじめとする顧客サービス等において頻繁に用いられている。

【0003】このような音声出力装置を用いる各種サービスでは、サービスの利用者は音声ガイダンスの送出の途中で音声による回答入力を行うことがある。これに

対して、多くの音声出力装置には、利用者が回答を行う際の発声の容易さを確保するために、その回答が開始された時点で利用者の発声の阻害要因となり得る音声ガイダンスの送出を停止する機能が付与されている。

【0004】

【発明が解決しようとする課題】しかしながら、このような音声出力装置が受信する信号は、回答のための音声ばかりではなく、咳払い、応答者の近傍に存在する扉の開閉音、自動車の走行音等の雑音を、回答のための音声であると誤認し、音声ガイダンスの送信を中断してしまう場合があり、利用者及び当業者に多大な負担を強いている。

【0005】このような事情に鑑み、本発明は、回答を行うための音声が入力された時のみ音声ガイダンスの送出を停止する音声出力装置、音声出力方法、及び音声出力プログラムを提供することを目的とする。

【0006】

【課題を解決するための手段】本発明の第 1 の特徴は、音声信号を送出する送出手段と、送出手段による音声信号の送出中に音声信号の受信先から入力された信号が音声であるか否かを判定し、入力された信号が音声であるときには音声信号の送出を停止させる送出停止手段とを有する音声出力装置であることを要旨とする。

【0007】上記の特徴を有する本発明にあつては、入力された信号が音声である場合にのみ音声ガイダンス等の音声信号の送出を停止し、音声以外の雑音等が入力された場合は、音声信号の送出を停止せずに続行する構成とした。これにより、雑音等の入力により音声信号の送出が誤って停止してしまうことを防止可能となる。さらに、音声信号の送出停止は、音声の入力が開始されると直ちに実行されるため、音声の入力者（発声者）は、音声信号に阻害されずに容易に入力（発声）を続行することが可能となる。

【0008】本発明の第 2 の特徴は、判定手段は、入力された信号の音圧レベルの測定及び入力された信号の基本周波数の算出を行う測定算出手段と、測定された音圧レベルと所定の閾値との比較及び算出された基本周波数の解析を行う比較解析手段と、測定した音圧レベルが閾値より大であり、且つ算出された基本周波数から入力された信号が音声であることを示す情報が検出された場合、入力された信号が音声であると判定する判定実行手段とを有することを要旨とする。

【0009】本発明の第 3 の特徴は、入力された信号を蓄積する入力信号蓄積手段を有することを要旨とする。

【0010】上記の特徴を有する本発明にあつては、判定部によって判定がなされた音声を蓄積する。また、雑音等を含む全ての信号を蓄積することもできる。これらの点により、更に充実したサービスを提供することが可能となる。

【0011】本発明の第 4 の特徴は、判定がなされた音

声を認識し、出力する音声認識出力手段を有することを要旨とする。

【0012】上記の特徴を有する本発明にあっては、判定手段によって判定がなされた音声を音声認識出力手段によって音声認識し、音声の発声内容を出力することが可能になる。判定手段により判定がなされた音声を音声認識出力手段へ入力とすることにより、音声認識出力手段は高い精度で発声内容を認識し出力することが可能になる。この点から、より充実したサービスを提供することが可能となる。なお、ここでいう音声の認識及び出力

には、音声の文字データへの変換、及び変換された文字データの出力も含まれる。

【0013】本発明の第5の特徴は、音声信号を送出する送出ステップと、送出ステップにおける音声信号の送出中に音声信号の受信先から入力された信号が音声であるか否かを判定し、入力された信号が音声であるときには音声信号の送出を停止させる送出停止ステップとを有する音声出力方法であることを要旨とする。

【0014】上記の特徴を有する本発明にあっては、入力された信号が音声である場合にのみ音声ガイダンス等の音声信号の送出を停止し、音声以外の雑音等が入力された場合は、音声信号の送出を停止せずに続行する。これにより、雑音等の入力により音声信号の送出が誤って停止してしまうことを防止可能となる。さらに、音声信号の送出停止は、音声の入力が開始されると直ちに実行されるため、音声の入力者（発声者）は、音声信号に障害されずに容易に入力（発声）を続行することが可能となる。

【0015】本発明の第6の特徴は、音声信号を送出する送出ステップと、送出ステップにおける音声信号の送出中に音声信号の受信先から入力された信号が音声であるか否かを判定し、入力された信号が音声であるときには音声信号の送出を停止させる送出停止ステップとを実行させる音声出力プログラムであることを要旨とする。

【0016】上記の特徴を有する本発明にあっては、入力された信号が音声である場合にのみ音声ガイダンス等の音声信号の送出を停止し、音声以外の雑音等が入力された場合は、音声信号の送出を停止せずに続行する。これにより、雑音等の入力により音声信号の送出が誤って停止してしまうことを防止可能となる。さらに、音声信号の送出停止は、音声の入力が開始されると直ちに実行されるため、音声の入力者（発声者）は、音声信号に障害されずに容易に入力（発声）を続行することが可能となる。

【0017】

【発明の実施の形態】以下、図面を用いつつ本発明の音声入力装置、音声入力方法、及び音声入力プログラムについての説明を行う。

【0018】なお、以下の実施の形態は、あくまでも本

発明の説明のためのものであり、本発明の特許請求の範囲を制限するものではない。したがって、当業者であれば、これらの各要素又は全要素を含んだ各種の実施の形態を採用することが可能であるが、これらの実施の形態も本発明の範囲に含まれる。また、実施の形態を説明するための全図において、同一要素のものは同一符号を付与し、これに関する反復説明は省略する。

【0019】図1は、本発明の第1の実施の形態にかかる音声出力装置の構成を示すブロック図である。音声出力装置1aは、記憶部2、請求項の送出手段に対応する送出部3、送出停止手段に対応する送出停止部4、入力信号蓄積手段に対応する入力信号蓄積部5、及び音声認識出力手段に対応する音声認識出力部6を有し、音声送信部2は音声出力端子7と接続し、送出停止部3、入力信号蓄積部5、及び音声認識出力部6は音声入力端子8と接続している。

【0020】音声出力端子7は、ネットワークを経由して通話装置のスピーカやハンドセットと接続し、音声入力端子7は、ネットワークを経由して通話装置のマイクやハンドセットと接続している。なお、ネットワークを介さずにスピーカ、マイク、ハンドセット等と接続する場合もある。

【0021】記憶部2は、音声信号を記憶する。なお、本実施の形態及び以降の全ての実施の形態においては、この音声信号が音声ガイダンスである例を示すが、これに限定されない。

【0022】送出部3は、音声ガイダンスの送出を要求する信号を受信すると、音声ガイダンスを記憶部2から読み出し、音声出力端子7を介して前述のスピーカ等に向けて送出する。

【0023】送出停止部4は、送出部3による音声ガイダンスの送出中に、前述のマイク等から音声入力端子8を介して何らかの信号が入力された際、その信号が音声であるときには、音声ガイダンスの送出を停止させる。なお、本発明における音声とは音声ガイダンスに対する回答の際に発せられる肉声等を指す。

【0024】このため、送出停止部4は、図2に示すとおり、入力された信号（以下、入力信号とする）が音声であるか否かの判定を行う判定部41（請求項の判定手段に対応）と停止信号送信部45とを有する。

【0025】この判定部41は、請求項の測定算出手段に対応する測定算出部42と、比較解析手段に対応する比較解析部43と、判定実行手段に対応する判定実行部44とを有する。

【0026】測定算出部42は、入力信号の音圧レベルの測定と、基本周波数の算出とを行う。比較解析部43は、測定算出部42により測定された音圧レベルと、所定の閾値とを比較し、同じく測定算出部42により算出された基本周波数の解析とを行う。

【0027】判定実行部44は、比較解析部43による

10

20

30

40

50

比較及び解析の結果、入力信号の音圧レベルが所定の閾値より大であり、且つ基本周波数の中から入力信号が音声であることを示す情報としてのピッチ成分が検出された場合、この入力信号が音声であると判定する。

【0028】停止信号送信部45は、判定実行部44による判定をもとに音声ガイダンスの送出停止信号を送出部3へ送信する。送出部3は、この送出停止信号を受信すると音声ガイダンスの送出を停止する。

【0029】一方、音圧レベルが所定の閾値より小である場合、またはピッチ成分が含まれていない場合は、音声ガイダンスの送出は停止されずに続行される。

【0030】音声ガイダンスは、音声の入力が終了するまで停止され、送出部3は、音声の入力が終了された旨を検知すると音声ガイダンスの送信を再開する。

【0031】なお、所定の閾値は任意の値に設定可能である。また、基本周波数は、声帯の振動周期に依存するが、一般的に男性で50～150ヘルツ、女性では200～300ヘルツの間の値を指す。

【0032】このようなピッチ成分の検出は、「音声波形の自己相関分析」、「ケプストラム分析」、「LPC残差信号分析」等の分析手法を用いることにより行われる。

【0033】なお、「音声波形の自己相関分析」の詳細については、伊達玄（訳）『デジタル信号処理（上、下）』（コロナ社、1978年）4章8節、及び古井貞熙『音響・音声工学』（近代科学社、1992年）11章3節を参照されたい。

【0034】また、「ケプストラム分析」の詳細については、『音響・音声工学』（前掲書）11章2節から3節、鹿野清宏『音声・音情報のデジタル信号処理』（昭晃堂、1997年）1章4節、及び齋藤収三『音声情報処理の基礎』（オーム社、1981年）7章3節を参照されたい。

【0035】また、「LPC残差信号分析」の詳細については、『音響・音声工学』（前掲書）11章4節、及び『音声情報処理の基礎』（前掲書）6章2節を参照されたい。

【0036】上記のような送出停止部4による入力信号の出力測定、基本周波数の算出、測定された音圧レベルと所定の閾値との比較、及びピッチ成分の検出は、送出部3による音声ガイダンスの送出中に入力される全ての信号に対して実行される。

【0037】入力信号蓄積部5は、入力信号の蓄積を行う。この入力信号蓄積部5は、判定部31によって判定がなされた音声のみを蓄積することの他、雑音等を含む全ての信号を蓄積することもできる。

【0038】音声認識出力部6は、判定がなされた音声を認識し、出力する。なお、ここでいう音声の認識及び出力には、音声の文字データへの変換、及び変換された文字データの出力も含まれる。

【0039】図3は、本発明の第2の実施の形態にかかる音声入出力装置の構成を示すブロック図である。本図においては、図1に示した音声出力端子7にスピーカ9が、音声入力端子8にマイク10が其々接続されている場合を示している。このような構成をとることにより、音声入出力装置1aを、街頭にて各種情報の提供・収集を行う街頭情報端末、各種サービスにおける自動応答受付装置等として用いることが可能となる。

【0040】なお、音声出力端子7及び音声入力端子8の記載は省略している。この点は以下に示す実施の形態においても同様である。

【0041】図4は、本発明の第3の実施の形態にかかる音声入出力装置の構成を示すブロック図である。本図においては、音声出力装置1aが電話網やインターネット網等の通信ネットワーク12を介して、スピーカ9とマイク10とを備える電話機やコンピュータ等の通話装置11と接続されている場合を示している。このような構成をとることにより、音声入出力装置1aを、前述の街頭情報端末や自動応答受付装置等の他に、インターネット網を介して提供されている各種のウェブ情報に音声を用いてアクセスするにあたってのポータル（玄関口）となるボイスポータルとして機能させることが可能となる。

【0042】図5は、本発明の第4の実施の形態にかかる音声入出力装置の構成を示すブロック図である。本図においては、図4におけるスピーカ9及びマイク10に代えて、これらの機能の両方を備えたハンドセット13を接続した場合を示している。このような構成をとることにより、音声入出力装置1aを、前述の街頭情報端末、自動応答受付装置、ボイスポータルに加えて、顧客からの苦情、要望等を自動的に受け付ける自動化コールセンタで用いることを可能とする。

【0043】図6は、本発明の第5の実施の形態にかかる音声入出力装置の構成を示すブロック図である。本図では、本発明の第1から第4の実施の形態にかかる音声入出力装置1aにエコーキャンセラ14を設けた音声入出力装置1bを示している。なお、音声入出力装置1bは、図5で示した場合と同様にハンドセット13を有する通話装置11と通信ネットワーク12を介して接続されている場合を示しているが、これに限定されず、図4で示したように、通信ネットワーク12を介さずにスピーカ9及びマイク10と接続することが可能である。

【0044】音声入出力装置1bの構成は、第1から第4の実施例において、音声出力端子7と音声入力端子8、スピーカ9とマイク10、あるいはハンドセット13のスピーカ部とマイク部との間にエコーが発生する経路であるエコーパスが存在する場合を想定している。

【0045】発声者の利用環境によっては、送出部3が送出した音声ガイダンスは、スピーカ9等から送出された後に空間中を伝播してマイク10に到達し、エコーと

して入力信号に重畳する。

【0046】このような環境下では音声ガイダンス送出中にエコーが入力されることがある。これらのエコーを除去するために前述のエコーキャンセラ14が設けられている。

【0047】従来の音声出力装置においては、エコーを消去する機能が付与されていてもエコーパスの変動やハードウェアの制約などから、エコーを完全に除去されず、残存したエコーが入力信号に重畳する場合があった。この残存エコーは前述のピッチ成分情報を含む場合があるが、一般にエコーの音圧レベルは小さい。

【0048】本実施の形態における音声出力装置1bは、第1から第4の実施の形態にかかる音声出力装置1aと同様に入力信号の音圧レベルを測定するため、所定の閾値をマイク10等から入力される残存エコーの音圧レベルの予想値より高い値に設定することにより、万が一残存エコーが入力信号に含まれても残存エコーを音声と誤認して音声ガイダンスの送出が停止することを防止可能である。

【0049】なお、第1から第5の実施の形態にかかる音声出力装置1a及び1bは、記憶部2を有する場合を示したが、これに限定されず、記憶部2を廃し、音声出力装置1a及び1bとは別個に設けられた記憶装置から音声ガイダンス等の音声信号を受信し、送出する構成とする他、同じく別個に設けられたスピーカ等から発せられる音声信号を受信し、送出する構成とすることも可能である。

【0050】上記の音声出力装置1a又は1bは、図7に例示するようなコンピュータ100によって実現することが可能である。コンピュータ100は、CPU101、このCPU101により制御されるHDD（ハードディスク・ドライブ）102、ROM103、RAM104、キーボード105、ディスプレイ106、FDD（フロッピディスク・ドライブ）107、通信部108等を有し、さらにテープ・レコーダ（図示せず）等を有することもできる。

【0051】このようなコンピュータ100は、図1に示した各部の機能を発現させるためのコンピュータ・プログラムを読み込み、実行することにより、音声出力装置1a又は1bとして機能する。このコンピュータ・プログラムは、フロッピディスク、CD-ROM等の各種情報記録媒体に記録可能である。また、コンピュータ100は、このようなコンピュータ・プログラムをインターネット等の各種通信ネットワークを介して入手することもできる。

【0052】したがって、上記のコンピュータ・プログラム、及びこのコンピュータ・プログラムを記録したプログラム媒体も本発明の範囲に含まれる。なお、本発明の音声出力装置を実現するには、様々な通信端末を用いることが可能であり、上記のコンピュータ100に限

定されない。

【0053】図8は、本発明の第1から第5の実施の形態にかかる音声出力装置の作用を示すフローチャートである。なお、本図においては、音声出力装置1a及び1bは、音声出力端子7及び音声入力端子8と接続されている場合を示しているが、前述のスピーカ9、マイク10、通話装置11、通信ネットワーク12と接続されている場合においても以下の記す音声出力装置1a及び1bが及ぼす作用に差異はない。

【0054】送出部3は、所定の音声ガイダンス送信開始信号を受信すると、記憶部2から音声ガイダンスを読み出し、音声出力端子7及びこれと接続されたスピーカ等に対して送信を開始する（ステップS1）。

【0055】ガイダンス送出部3による音声ガイダンスの送出中に、送出停止部4に対して前述のマイク等から音声入力端子8を介して、何らかの信号が入力されると、この送出停止部4の判定部41は、入力信号の音圧レベルを測定し（ステップS2）、この音圧レベルと所定の閾値とを比較し（ステップS3）、音圧レベルが所定の閾値より小であれば（ステップS3；NO）、音声ガイダンスの送出が続行される。

【0056】一方、音圧レベルが所定の閾値より大であれば（ステップS3；YES）、この入力信号の基本周波数が算出され（ステップS4）、この基本周波数の解析つまり入力信号が音声であることを示唆するピッチ成分の検出が行われる（ステップS5）。

【0057】解析の結果、このピッチ成分が検出されなかった場合（ステップS5；NO）、音声ガイダンスの送出は続行される。

【0058】一方、ピッチ成分が検出された場合（ステップS5；YES）、入力信号が音声であると判定され、同じく送出停止部4が有する停止信号送信部45は、音声ガイダンスの送出停止信号を送出部3へ送信する（ステップS6）。

【0059】送出停止信号を受信したガイダンス送出部3は、音声ガイダンスの送出を停止する。

【0060】音声ガイダンスは、音声の入力が終了するまで停止され、ガイダンス送出部3は、音声信号の入力が終了した旨を検知すると音声ガイダンスの送信を再開する。この音声信号の入力が終了した旨は、発話者の沈黙、所定のプッシュ・ボタンの押下等により生ずる信号終了信号により示される。

【0061】以上のような動作は、音声ガイダンスの送出が終了するまで全ての入力信号に対して行われる。また、必要に応じて図1、及び図3から図7の入力信号蓄積部5による入力信号蓄積ステップと音声認識出力部6による音声識別が行われる（共に図示せず）。

【0062】なお、本図においては、音声出力装置1a及び1bの動作は、音声入力信号の出力測定（ステップS2）、音圧レベルと所定の閾値との比較（ステップ

10

20

30

40

50

S3)、基本周波数の算出(ステップS4)、入力信号が音声であるか否かの判定(ステップS5)の順で行われる場合を示したが、ステップS2とステップS4とを実行した後ステップS3とステップS5を実行する構成とすることも可能である。

【0063】

【発明の効果】以上のとおり本発明は、音声信号、つまり回答を行うための音声が入力された時にのみ音声ガイダンスの送出を停止する構成とした。これにより、雑音を、回答のための音声であると誤認し、音声ガイダンスの送出を中断してしまうことを防止可能となった。

【0064】さらに、万が一のエコーを完全に消去できなかった際にもその音圧レベルを所定の閾値と比較する構成としたことにより、エコーによる音声ガイダンスの送出停止を防止可能となった。このような点から、発声者が快適且つ簡易に音声入力を行える音声出力装置を提供することが可能となった。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態にかかる音声出力装置の構成を示すブロック図である。

【図2】図1に示す送出停止部の詳細を示すブロック図である。

【図3】本発明の第2の実施の形態にかかる音声出力装置の構成を示すブロック図である。

【図4】本発明の第3の実施の形態にかかる音声出力装置の構成を示すブロック図である。

【図5】本発明の第4の実施の形態にかかる音声出力装置の構成を示すブロック図である。

【図6】本発明の第5の実施の形態にかかる音声出力装置の構成を示すブロック図である。

【図7】図1から図6に示す音声出力装置として用いられるコンピュータの構成の一例を示すブロック図であ*

る。

【図8】図1から図6に示す音声出力装置の作用を示すフローチャートである。

【符号の説明】

1a、1b 音声出力装置

2 記憶部

3 送出部

4 送出停止部

5 入力信号蓄積部

6 音声認識出力部

7 音声出力端子

8 音声入力端子

9 スピーカ

10 マイク

11 通話装置

12 通信ネットワーク

13 ハンドセット

14 エコーキャンセラ

41 判定部

42 測定算出部

43 比較解析部

44 判定実行部

45 停止信号送信部

100 コンピュータ

101 CPU

102 HDD

103 ROM

104 RAM

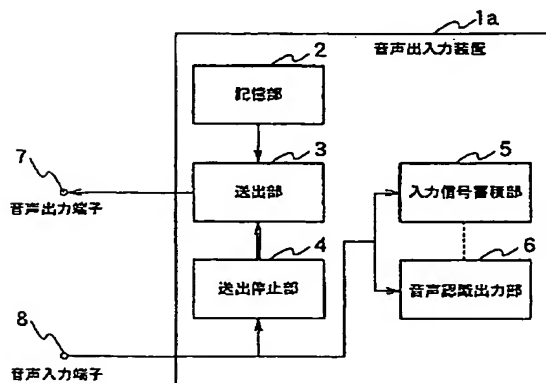
105 キーボード

106 ディスプレイ

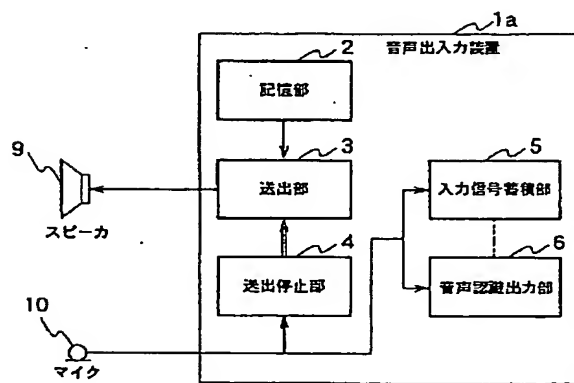
107 FDD

108 通信部

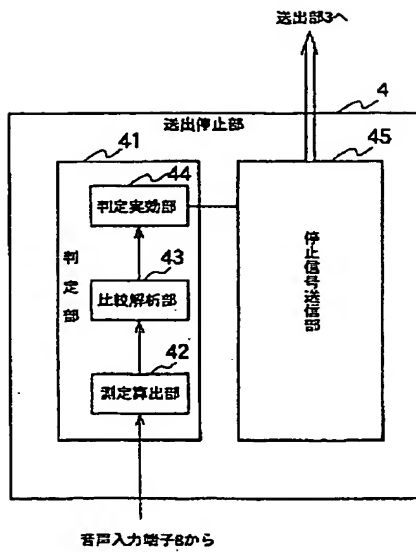
【図1】



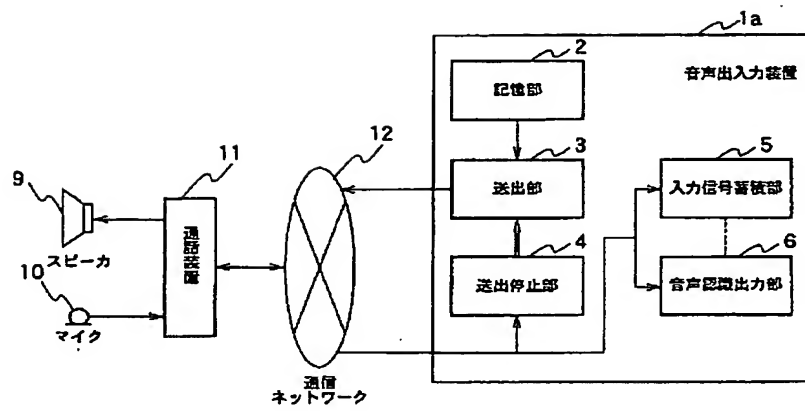
【図3】



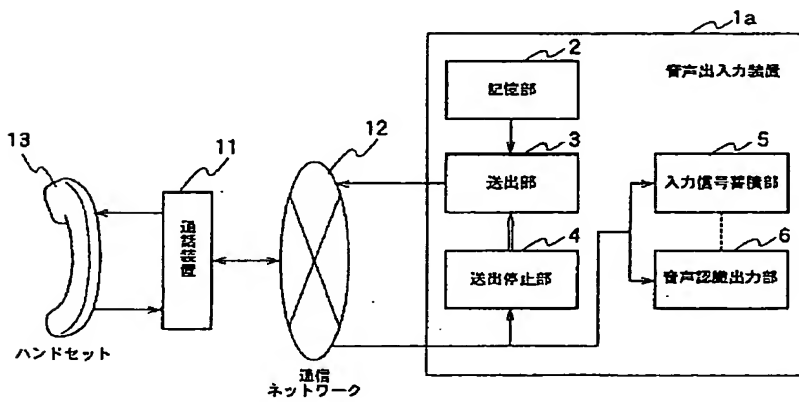
【図2】



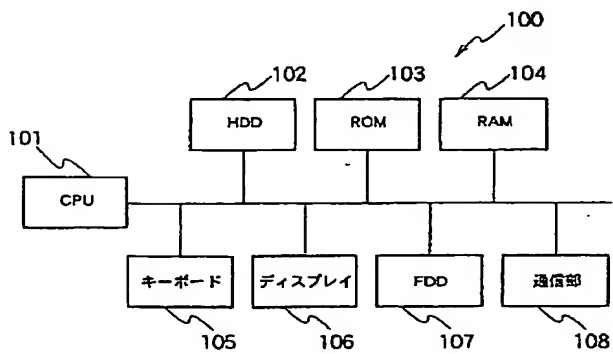
【図4】



【図5】



【図7】



The diagram illustrates a speech signal processing system. At the top, four main components are shown: 7 (音声出力端子 - Voice Output Terminal), 8 (音声入力端子 - Voice Input Terminal), 3 (送出部 - Output Section), and 4 (送出停止部 - Output Stop Section). Below these, a flowchart details the logic. It starts with '音声ガイダンス' (Voice Guidance) being sent to the output terminal (7) and the input terminal (8). An '入力信号' (Input Signal) is received at the input terminal (8) and sent to the output section (3). The output section (3) sends '音声ガイダンス送出' (Voice Guidance Output, S1) to the output terminal (7). The input signal is also sent to the output stop section (4). The output stop section (4) performs '入力信号出力判定' (Input Signal Output Determination, S2). A decision is made at S3: '音圧レベルは閾値より大きいのか?' (Is the sound pressure level greater than the threshold?). If YES, it proceeds to S4: '入力信号の基本周波数算出' (Calculation of the fundamental frequency of the input signal). If NO, it proceeds to S5: '音声ガイダンス送出続行' (Continue voice guidance output). From S4, a decision is made at S5: '音声信号とあることと一致しなかったか?' (Did the voice signal not match?). If YES, it proceeds to S6: '送出停止信号送信' (Transmit output stop signal). If NO, it proceeds to S7: '音声ガイダンス送出停止' (Stop voice guidance output). The output stop section (4) sends the '送出停止信号' (Output stop signal) to the output section (3), which then stops the voice guidance output (S7).

```

graph TD
    7[音声出力端子]
    8[音声入力端子]
    3[送出部]
    4[送出停止部]
    
    8 -- "音声ガイダンス" --> 7
    8 -- "音声ガイダンス" --> 8
    8 -- "入力信号" --> 3
    3 -- "音声ガイダンス送出 S1" --> 7
    
    8 -- "入力信号" --> 4
    
    4 -- "入力信号出力判定 S2" --> S3{音圧レベル  
は閾値より大きいのか?}
    
    S3 -- NO --> S5{音声信号と  
あることと一致し  
なかったか?}
    S3 -- YES --> S4[入力信号の  
基本周波数算出]
    
    S4 --> S5
    
    S5 -- YES --> S6[送出停止  
信号送信]
    S5 -- NO --> S7[音声ガイダンス  
送出停止]
    
    S6 -- "送出停止信号" --> 3
    3 -- "音声ガイダンス送出停止 S7" --> 7
  
```

F ターム(参考) 5D015 CC01 CC14 DD00 FF00 LL05